Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 41, pp. 338-341, May 2025.

A NOTE ON LIDSKII'S THEOREM FOR THE DIRECTIONAL DERIVATIVES OF THE EIGENVALUES OF SYMMETRIC MATRICES*

HRISTO SENDOV†

Abstract. Lidskii's theorem is a classical perturbation result for the spectrum of symmetric matrices. It states that a change in the spectrum caused by a perturbation is majorized by the original spectrum. We present a generalization and a refinement of Lidskii's theorem involving the directional derivatives of the spectrum of a symmetric matrix.

Key words. Lidskii's theorem, Schur's theorem, Eigenvalues, Directional derivative, Majorization.

AMS subject classifications. 47A55, 15A18, 90C22.

1. Majorization. Denote by P^n the set of all $n \times n$ permutation matrices. For any vector $x \in R$ denote by $[x]$ the vector obtained from x by permuting its coordinates in nonincreasing order $[x]_1 \geq \cdots \geq [x]_n$. Denote by $\mathbb{R}^n_{\geq}$ the set of all vectors in $\mathbb{R}^n$ with coordinates ordered non-increasingly.

Definition 1. For any two vectors $x, y \in \mathbb{R}^n$, we say that *x is majorized by y*, denoted by $x \prec y$, if the following relationships hold

$$
\begin{aligned}
[x]_1 &\leq [y]_1, \\
[x]_1 + [x]_2 &\leq [y]_1 + [y]_2, \\
&\vdots \\
[x]_1 + \cdots + [x]_{n-1} &\leq [y]_1 + \cdots + [y]_{n-1}, \\
[x]_1 + \cdots + [x]_{n-1} + [x]_n &= [y]_1 + \cdots + [y]_{n-1} + [y]_n.
\end{aligned}
$$

The following famous theorem is due to Weyl, Birkhoff, and Hardy-Littlewood-Polya (see [2], [4], [6], [7]).

Theorem 2. *Let $x, y \in \mathbb{R}^n$. Then the following conditions are equivalent:*

1. $x \prec y$.
2. $x \in \operatorname{conv} \{Py \,|\, P \in P^n\}$.
3. $a^T x \leq [a]^T [y]$ *for all* $a \in \mathbb{R}^n$.
4. *There exists an $n \times n$ doubly stochastic matrix S, such that $x = Sy$.*
5. *The inequality*
$$\sum_{i=1}^{n} f(x_i) \leq \sum_{i=1}^{n} f(y_i),$$
holds for any convex function f on $\mathbb{R}$.

*Received by the editors on April 9, 2025. Accepted for publication on May 6, 2025. Handling Editor: Ren-Cang Li. Corresponding Author: Hristo Sendov.

†Department of Statistical and Actuarial Sciences, Department of Mathematics, Western University, 1151 Richmond Str., London, ON, N6A 5B7 Canada, (hsendov@uwo.ca). The author was partially supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada. (Grant number RGPIN-2020-06425.)

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 41, pp. 338-341, May 2025.

6. *The inequality*

$$f(x_1,\ldots,x_n) \le f(y_1,\ldots,y_n),$$

holds for any convex symmetric function f on R^n.

We now introduce a refinement of the notion of majorization. Let π be a partitioning of the set $\mathbb{N}_n := \{1,2,\ldots,n\}$. Each element of the partitioning is called a *block*. Denote by P^n_π the subgroup of P^n that preserves the blocks of the partitioning.

Definition 3. For any two vectors $x, y \in \mathbb{R}^n$, we say that *x is π-majorized by y*, denoted by $x \prec_\pi y$, if $x \in \operatorname{conv}\{Py \mid P \in P^n_\pi\}$.

For example, when $\pi = \{\{1,2,\ldots,n\}\}$, then $x \prec_\pi y$ if and only if $x \prec y$. When $\pi = \{\{1\},\{2\},\ldots,\{n\}\}$, then $x \prec_\pi y$ if and only if $x = y$. Clearly, if $x \prec_\pi y$ for some partition π of $\mathbb{N}_n$, then $x \prec y$. Every vector λ in $\mathbb{R}^n$ defines a partition of the set $\mathbb{N}_n$ such that, two numbers i and j are in the same block if $\lambda_i = \lambda_j$. The partition defined in this way by a vector $\lambda \in \mathbb{R}^n$ is denoted by $\pi(\lambda)$.

2. Lidskii and Schur's theorems.

2.1. A generalization of Lidskii's theorem. Denote by S^n the space of $n \times n$ real symmetric matrices. For $A \in X$ denote the vector of eigenvalues, counting multiplicities by $\lambda(A) := (\lambda_1(A),\ldots,\lambda_n(A))$. We assume that the eigenvalues are ordered non-increasingly:

$$\lambda_1(A) \ge \cdots \ge \lambda_n(A).$$

One of the central results in matrix perturbation theory and a main theme in [1] is the following theorem.

Theorem 4 (Lidskii). *For any two matrices $A, B \in S^n$*

$$\lambda(A+B) - \lambda(B) \prec \lambda(A).$$

An immediate corollary of Lidskii's theorem and the definition of majorization is that the sum of the k largest eigenvalues of $A \in S^n$, denoted by $\sigma_k(A)$, is a convex function on S^n. Indeed,

$$\sum_{i=1}^{k} \big(\lambda_i(A+B) - \lambda_i(B)\big) \le \sum_{i=1}^{k} [\lambda(A+B) - \lambda(B)]_i \le \sum_{i=1}^{k} \lambda_i(A)$$

shows that the sum of the k largest eigenvalues is a subadditive function. Since the eigenvalues are also positively homogeneous, the convexity claim follows. Every convex function is directionally differentiable in the interior of their domain, [8]. Since $\lambda_i(A) = \sigma_i(A) - \sigma_{i-1}(A)$, we obtain that each eigenvalue is directionally differentiable. In other words, the following limit exists and is finite for every $A, X \in S^n$:

$$\lambda_i'(A;X) := \lim_{t \to 0^+} \frac{\lambda_i(A+tX) - \lambda_i(A)}{t}.$$

Several trivial examples are

$$\lambda'(A;I) = (1,\ldots,1)^T,\ \lambda'(I;A) = \lambda(A),\ \text{and}\ \lambda'(A;A) = \lambda(A).$$

Notice that $\lambda_i'(\alpha A;\beta X) = \beta\lambda'(A;X)$ for all $\alpha,\beta \ge 0$. Our goal is to show the following generalization of Lidskii's theorem.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 41, pp. 338-341, May 2025.

THEOREM 5 (Lidskii for directional derivatives). *For any $A, X, Y \in S^n$*

$$\lambda'(A; X+Y) - \lambda'(A; Y) \prec_{\pi(\lambda(A))} \lambda'(A; X). \tag{2.1}$$

The statement is a genuine generalization of Lidskii's theorem, since if we take $A = I$ in (2.1), we obtain $\lambda(X+Y) - \lambda(Y) \prec \lambda(X)$, using that $\pi(\lambda(I)) = \{\{1, 2, \dots, n\}\}$. On the other hand, Theorem 5 is not a trivial consequence of Theorem 4. Indeed, observe first that Theorem 4 implies

$$\frac{\lambda(A + t(X+Y)) - \lambda(A + tY)}{t} \in \operatorname{conv}\{P\lambda(X) \,|\, P \in P^n\}, \quad \text{for all } t > 0.$$

Adding and subtracting $\lambda(A)$ in the numerator and taking the limit as $t \to 0^+$ give

$$\lambda'(A; X+Y) - \lambda'(A; Y) \prec \lambda(X). \tag{2.2}$$

Analogously, from

$$\frac{\lambda(A + tX) - \lambda(A)}{t} \in \operatorname{conv}\{P\lambda(X) \,|\, P \in P^n\}, \quad \text{for all } t > 0,$$

in the limit as $t \to 0^+$, we obtain

$$\lambda'(A; X) \prec \lambda(X). \tag{2.3}$$

Thus, (2.2) and (2.3) together are weaker statements than (2.1).

2.2. Schur's theorem for directional derivatives. A theorem that is closely related to Lidskii's theorem is the Schur's theorem. Define the linear map diag : $S^n \to \mathbb{R}^n$ by $\operatorname{diag}(A) := (A_{11}, A_{22}, \dots, A_{nn})^T$, the vector of diagonal entries of A. Its adjoint map is given by Diag : $\mathbb{R}^n \to S^n$, where $\operatorname{Diag}(a)$ is the diagonal matrix with vector a on its diagonal.

THEOREM 6 (Schur). *For any symmetric matrix $A \in S^n$*

$$\operatorname{diag}(A) \prec \lambda(A).$$

A generalization of Schur's theorem for the directional derivatives of the eigenvalues was proved in [5, Theorem 4]. The proof there is involved.

THEOREM 7 (Schur for directional derivatives). *Any vector $a \in \mathbb{R}^n_{\geq}$ and a matrix $X \in S^n$ satisfy*

$$\operatorname{diag}(X) \prec_{\pi(a)} \lambda'(\operatorname{Diag} a; X). \tag{2.4}$$

Taking vector a to be the all one vector, we get $\lambda'(\operatorname{Diag} a; X) = \lambda(X)$, showing that Theorem 7 is indeed a generalization of Theorem 6. Combining (2.3) and (2.6) shows that Schur's theorem is a limiting case of Lindskii theorem. Finally, (2.3) gives the following refinement

$$\operatorname{diag}(X) \prec_{\pi(a)} \lambda'(\operatorname{Diag} a; X) \prec \lambda(X).$$

2.3. Proof of Theorem 5. As pointed out, the eigenvalues are directionally differentiable as a function of the symmetric matrix argument. A formula for the directional derivative was first obtained in [3, Theorem 3.12]. We now describe it.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 41, pp. 338-341, May 2025.

Let $A, X \in S^n$. Assume that A has r distinct eigenvalues with the following multiplicities:

$$\lambda_1(A) = \cdots = \lambda_{k_1}(A) > \lambda_{k_1+1}(A) = \cdots = \lambda_{k_2}(A) \\ > \cdots > \lambda_{k_{r-1}+1}(A) = \cdots = \lambda_{k_r}(A).$$

It is convenient to define $k_0 := 0$. Let U be an orthogonal matrix such that $A = U(\operatorname{Diag}\lambda(A))U^T$ is the ordered spectral decomposition of A. Let $e_1, \ldots, e_n$ be the standard basis in $\mathbb{R}^n$ and let u_i be the ith column of U. For every $i = 1, \ldots, r$ define the $n \times (k_i - k_{i-1})$ matrices

$$U_i := [u_{k_{i-1}+1}, \ldots, u_{k_i}] \text{ and } E_i := [e_{k_{i-1}+1}, \ldots, e_{k_i}].$$

Then

$$\begin{aligned} \lambda'(A; X) &= (\lambda(E_1^T U X U^T E_1)^T, \lambda(E_2^T U X U^T E_2)^T, \ldots, \lambda(E_r^T U X U^T E_r)^T)^T \\ &= (\lambda(U_1^T X U_1)^T, \ldots, \lambda(U_r^T X U_r)^T)^T. \end{aligned} \tag{2.5}$$

If A has distinct eigenvalues, then we simply have

$$\lambda'(A; X) = \operatorname{diag}(U^T X U),$$

and in this case the eigenvalues are analytic functions locally around A. In particular, if vector $a \in \mathbb{R}^n_{\geq}$ has distinct entries, then

$$\lambda'(\operatorname{Diag} a; X) = \operatorname{diag}(X). \tag{2.6}$$

Using Formula (2.5) and the classical Lidskii's theorem applied to each block gives an immediate verification of Theorem 5.

The derivation of Formula (2.5) is quite involved and that motivates a search for a direct approach to Theorem 5. That, most likely, was the motivation behind the direct proof of Theorem 7 given in [5].

REFERENCES

[1] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.
[2] G. Hardy, J.E. Littlewood, and G. Polya. *Inequalities*, 2nd edition. Cambridge University Press, Cambridge, 1989.
[3] J.-B. Hiriart-Urruty and D. Ye. Sensitivity analysis of all eigenvalues of a symmetric matrix. *Numer. Math.*, 70:45–72, 1995.
[4] L. Hörmander. *Notions of Convexity*. Progress in Mathematics, vol. 127. Birkhäuser, Boston, 1994.
[5] A.S. Lewis. Nonsmooth analysis of eigenvalues. *Math. Program.*, 84:1–24, 1999.
[6] M. Marcus and H. Minc. *A Survey on Matrix Theory and Matrix Inequalities*. Allyn and Bacon, Boston, 1964.
[7] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Mathematics in Science and Engineering, vol. 143. Academic Press, New York, 1979.
[8] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.